

Research Article

Balance Assessment in Multiple Sclerosis and Cerebellar Ataxia: Rationale, Protocol and Demographic Data

Winser S^{1*}, Smith CM¹, Hale LA¹, Claydon LS², Whitney S^{3,4}, Klatt B³, Mottershead J⁵, Zaydan I⁶ and Heyman R⁶

¹Centre for Health, Activity and Rehabilitation Research, School of Physiotherapy, University of Otago, New Zealand

²Department of Allied Health and Medicine, Anglia Ruskin University, UK

³School of Health and Rehabilitation Sciences, University of Pittsburgh, USA

⁴Rehabilitation Research Chair at King Saud University, Saudi Arabia

⁵Dunedin School of Medicine, Neurology, University of Otago, New Zealand

⁶School of Medicine, Neurology, University of Pittsburgh, USA

*Corresponding author: Stanley Winser, Centre for Health, Activity and Rehabilitation Research, School of Physiotherapy, University of Otago, New Zealand

Received: November 11, 2014; Accepted: November 20, 2014; Published: November 25, 2014

Abstract

A core set of standardized balance measures are required for use in rehabilitation among people with multiple sclerosis (MS) and cerebellar ataxia. An earlier systematic review and Delphi survey identified the Berg Balance Scale (BBS), the Timed Up and Go test (TUG), Posture and Gait sub-component of the International Co-operative Ataxia Rating Scale (PG of ICARS) and the gait, sitting and stance sub-components of the Scale for the Assessment and Rating of Ataxia (SARA Bal) as suitable balance measures. This study aims to estimate the reliability, validity and interpretability of these measures. This study will recruit 60 participants with multiple sclerosis with secondary cerebellar involvement across four centres in New Zealand and the United States of America. Participants will be assessed and videotaped performing the BBS, TUG, SARA Bal and PG of ICARS by trained physiotherapists. Barthel Index, Expanded Disability Status Scale (EDSS), Disease duration, ICARS and SARA will also be assessed to determine validity. A second assessment to determine reliability will be conducted by assessors watching the video-recording. Data collection is in progress, 44 samples have been collected and the demographic data are presented. The findings of this study will recommend a core set of reliable, valid and interpretable measures that are suitable for clinical practice and research for the assessment of balance among adults with MS and cerebellar ataxia. Minimal Clinically Important Difference (MCID) and cut-off scores to predict the use of assistive walking device will be established.

Keywords: Study protocol; Reliability; Validity; Multiple sclerosis and Cerebellar ataxia

Abbreviations

BBS: Berg Balance Scale; COSMIN: Consensus based Standards for the selection of health status Measurement INstrument; EDSS: Expanded Disability Status Scale; ICC: Intra class Correlation Coefficient; MCID: Minimally Clinically Important Difference; MS: Multiple Sclerosis; NZ: New Zealand; PG of ICARS- Posture and Gait sub-component of the International Co-operative Ataxia Rating Scale; SARA Bal: gait, stance and sit sub-components of the Scale for the Assessment and Rating of Ataxia; TUG: Timed Up and Go test; USA: United State of America; α : Cronbach's alpha.

Introduction

Multiple sclerosis (MS) is a progressive, demyelinating disorder of the central nervous system resulting in a wide spectrum of clinical symptoms that include physical and cognitive impairments [1]. MS has a global prevalence ranging between 200 per 100,000 to very rare cases in some countries [2, 3]. New Zealand (NZ) has an overall prevalence of 73.1/100,000 of people diagnosed with MS, and that prevalence increases three-fold in the South Island [4] raising to 134.6/100,000 [5]. The United States of America (USA) has a prevalence of 90 / 100,000 [6]. Up to 80% of people with MS experience ataxia at some point in the disease course [7]. Demyelination due to MS involves multiple areas of the brain and it may be difficult to localize the lesion based on the presenting symptoms. However, ataxia, limb dyskinesia and postural tremor occurring in MS are believed to be

due to lesions of the cerebellum and its connections [8]. Cerebellar ataxia secondary to MS presents with one or more of the following symptoms: ataxia (both limb and gait), nystagmus, disequilibrium, dysarthria and hypotonia [9].

Cerebellar ataxia itself is not a disease but an 'umbrella' term that includes multiple diseases with lesions localized to the cerebellum. Choosing an appropriate measure for the assessment of balance in cerebellar ataxia is challenging as there is no recommended set of tools, yet a wide range of measures are currently used for interventions and identification or illustration of clinical features relating to balance and postural control [10]. Current guidelines focus on treatment strategies to improve balance and gait in people with cerebellar ataxia, and little emphasis has been placed on the optimal use of measures of balance [7, 11, 12]; only one international guideline [13] reports on best clinical practice for the intervention and assessment of balance in people with cerebellar ataxia and present a list of 13 measures of balance that are commonly used in clinical practice. The diversity of measures of balance used in published information is not helpful to derive a core set of standardized measures of balance among people with cerebellar ataxia. It would be useful to determine how many of the measures used in literature are appropriate for the assessment of balance among the target group. Interestingly, a systematic review that evaluated the effects of physiotherapy treatment options for cerebellar ataxia reports the lack of use of valid and reliable outcome measures to estimate treatment benefits [14].

Frequently used balance measures in the clinical setting for people with cerebellar ataxia are: Single leg stance [15], Tandem standing [16], Standing with eyes closed (Romberg's test) [17], Walking in a straight line (Tandem Walking), Walking along around a circle, Functional Reach Test [18] and External perturbation test [15]. Although these functional measures provide useful information on changes in balance, their psychometric properties and appropriateness for the assessment of balance is not known.

Earlier research to identify a core set of standardized measures of balance among people with cerebellar ataxia comprised a systematic review and a Delphi survey. The systematic review identified the Posture and gait sub-component of the International Co-operative Ataxia Rating Scale (PG of ICARS) as an appropriate measure because of its favorable psychometric properties [19]. The Delphi survey asked neurologists and physiotherapists specialized in treating people with cerebellar ataxia about clinically useful measures. The outcomes of the survey indicated the Berg Balance Scale (BBS), Timed up and go (TUG) test and the gait, stance and sit sub-components of the Scale for the Assessment and Rating of Ataxia (SARA Bal) as recommended measures of balance [20]. In order to strengthen recommendations for use of these four measures of balance in clinical practice, further validation is required. Due to the high prevalence of MS in NZ and USA, people with MS and secondary cerebellar involvement will be recruited and tested for this validation.

The aims of this study are to estimate the reliability, validity and interpretability of the BBS, TUG, PG of ICARS and SARA Bal among people with MS and secondary cerebellar ataxia. The specific aims of this study are (1) To estimate the internal consistency, intra-rater and inter-rater reliability of the four measures of balance, (2) To estimate the criterion, convergent, external and predictive validity of the measures of balance, (3) Determine the cut-off scores, sensitivity and specificity for the measures to discriminate between assistive walking device users and non-users and (4) To derive the Minimally Clinical Important Difference (MCID) for the four measures of balance.

Materials and Methods

The study is registered with the Australia New Zealand Clinical Trials Registry (Ref: ACTRN12613000079741) and ethical clearance has been obtained from the University of Otago Human Ethics Committee (Ref: 13/041), NZ and Institutional Review Board (IRB), The Multiple Sclerosis and Related Disorders Research Registry (Ref: PRO12010609) USA.

A cross sectional methodological observation study is proposed. As recommended by the CONsensus based Standards for the selection of health status Measurement INSTRument (COSMIN) checklist for 'good' quality research, a sample size of 50 is required [21] and this study aims for a sample size of 60. The data will be collected in three centres across NZ and one centre in the USA. People with MS and secondary cerebellar involvement will be recruited if they are: aged between 18 and 65 years, people with a definite diagnosis of MS presenting with at least one of the following clinical cerebellar symptoms: gait ataxia, limb ataxia - identified by dysdiadochokinesia and dysmetria, dysarthria, nystagmus and they are able to walk at least 10m with or without the use of assistive aids. Exclusion criteria include people with inability to follow simple commands, severe visual impairment, Expanded Disability Status Scale (EDSS) score >

6.5 and those who do not give permission for the research team to access to their medical record.

Procedure

The investigators will screen the volunteer for eligibility. All participants will be assessed on one occasion that will last for 60 minutes. The following measures will be assessed during this visit: (1) The BBS, (2) The TUG, (3) PG of ICARS and (4) the SARA Bal. The assessment process will be video-recorded (detailed below). The order of balance assessment tests is kept constant as follows: BBS, PG of ICARS, TUG and SARA Bal. The participants are split equally and the order of assessment is reversed (SARA Bal to BBS), in order to neutralise the influence of fatigue on the measure that is assessed last during the session. In addition, the Barthel Index, and full scales of ICARS and SARA will be scored by the investigators during the assessment session. These scales will be used to derive the constructs of validity. The EDSS score will be either retrieved from the participant's medical record or a separate appointment is arranged with the neurologist for this evaluation.

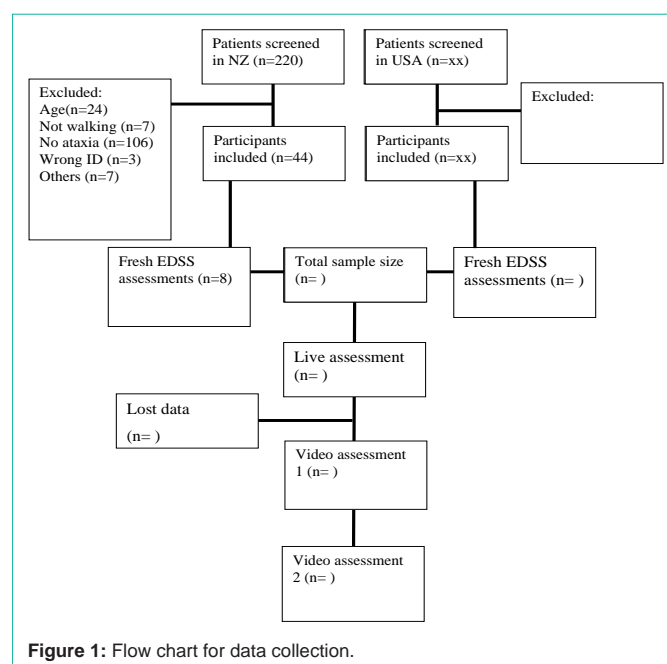
A video-recording of the four measures of balance (BBS, TUG, PG of ICARS and SARA Bal) will be performed using a wide angle digital video for each participant during the appointment. Sony 3.3 Mega Pixels handycam is being used for the video recording, in order to enhance a wide angle capture a Sony wide end conversion lens $\times 0.8$ is fitted to the camera lens. Research assistants are appointed to assess and video shoot the proceedings in both countries. To standardize the video recording across the study centres, a 'Standard video recording protocol' was developed by the research team and distributed among the investigators. The following points are emphasised in the 'Standard video recording protocol': the order of measures, instructions to the participants, equipment used for the assessment (thickness of mattress, height of chair, height of foot stool, use of footwear of the participant etc.), angle(s) of video recording, and the lighting of the assessment area. The 'Standard video recording protocol' minimises the potential inconsistencies in the above-mentioned factors that are crucial for obtaining an accurate reliability and validity estimate. After recording a test session, the video data will then be transcribed onto a DVD and distributed among the other members of the research team. The data from the USA will be transferred to NZ through Secure Zip, a password-enabled data transfer software. The video recordings will be used in this study to evaluate constructs of reliability.

Below, the first assessment done by the investigators on participants is called the 'Live assessment', and the repeat assessment done by same investigator from the video recording is called 'Video assessment 1'. The repeat assessment done by the second investigator from the same video recording is identified as 'Video assessment 2'. The flow of this study is outlined in Figure 1.

Psychometric properties tested

(Definitions adopted with permission from COSMIN manual [22].)

Intra-rater reliability: Defined as the proportion of variation in the scores of the participant done by the same assessor with an interval of 7-10 days. Variation between the live assessment scores and the scores of video assessment 1 will be compared. Video assessment 1 is the repeat assessment done by the same rater observing the video recording 7-10 days later.



Inter-rater reliability: Defined as the proportion of variation in the scores of the participant done by two different assessors. Variation between the scores of the live assessment and the scores of video recording 2 will be compared. The second rater will assess by observing the video recording.

Internal consistency: Defined as the degree of interrelatedness between the test items, internal consistency will be calculated for each measure separately.

Criterion validity: This is defined as the degree to which the scores of the measure under investigation are an adequate reflection of a 'gold standard'. Since a 'gold standard' is currently not available for balance assessment for people with MS with cerebellar ataxia, or those with other types of cerebellar ataxia, the measures will be correlated between each other.

Hypothesis testing: Defined as the degree to which the scores of the measures under investigation are consistent with the hypotheses (for instance, with regard to internal relationships, relationships to scores of other instruments, or differences between relevant groups), this is based on the assumption that the instrument validly measures the construct to be measured. The following hypotheses will be tested:

- Correlation between balance measures and ataxia rating scales (ICARS and SARA) - Convergent validity
- Correlation between balance measures and MS disease staging score (EDSS) - External validity
- Correlation between balance measures and functional independence (Barthel Index) - External validity

Predictive validity: Defined as the ability of the balance measures to differentiate between assistive walking device users and non-users, this will be determined by observing the group difference for the measures of balance. The cut-off score, sensitivity, and specificity of the measures of balance to discriminate between the groups will be identified.

Interpretability

MCID: This may be defined as the smallest difference a patient would perceive as beneficial as a result of the treatment [23]. The MCID will be established using the data driven method.

Measures considered

The BBS is a five point ordinal scale with 14 tasks that are each scored between 0 and 4, with the highest score of 56. This measure is interpreted as better balance with higher scores [24]. This measure has good inter-rater (ICC=0.96) and test retest (ICC=0.94) reliability [25, 26]. Acceptable concurrent validity in assessing balance among people with MS has been reported [27].

The TUG is a measure of an individual's dynamic stability and can predict falls risk. The TUG measures the time taken in seconds to arise from a chair, walk 3 meters, turn through 180 degrees and return to the seat [28]. A longer completion time indicates a higher risk of falling. The TUG has good test retest reliability (ICC=0.94) [26], moderate correlation with other clinical measures of balance, and an acceptable concurrent validity in assessing balance among people with MS [27].

The SARA is an ataxia rating scale. It consists of eight items among which the gait, sitting, and standing subcomponents are related to balance assessment. The three sub-components are scored out of 18 and called SARA Bal. The higher the score obtained, more severe the ataxia. SARA has high internal consistency ($\alpha > 0.85$) among the test items [29], very high inter rater (ICC>0.80) [30], intra rater (ICC>0.86) [31] and test retest reliability (ICC>0.91) [29]. Structural validity has been reported [32], satisfactory convergent validity when correlated with other ataxia rating scales [32], and adequate responsiveness has been demonstrated [33]. These tests have been conducted among both genetic and acquired forms of cerebellar diseases.

ICARS is a measure of ataxia severity and consists of 19 items categorised as 1: Posture and gait disturbances; 2: Kinetic function; 3: Speech disorders and 4: Oculomotor disorders. The posture and gait sub-component relates to balance assessment. A participant can score a maximum of 100, with a higher score denoting more severe ataxia. The posture and gait sub-component is scored out of 34. The ICARS has excellent intra rater (ICC=0.96) and test retest (ICC=0.96) [34] reliability and high internal consistency ($\alpha = 0.93$). Structural validity has been estimated [35]. The measure has adequate criterion validity [36] and good responsiveness [37]. The psychometric property testing has been used in a wide spectrum of diseases with cerebellar involvement.

The Barthel Index is an ordinal scale used to measure performance in activities of daily living (ADL). This scale rates the performance of ten items relating to ADL and mobility. Each item is scored between 0 and 15. The higher the score obtained, the better the functional independence. The scale has a total of 100 points and a minimum of 0. The Barthel Index has adequate to excellent inter-rater reliability (kappa values between 0.53 and 0.94) and excellent internal consistency ($\alpha = 0.89$) among people with acute stroke [38]. The scale has excellent criterion [39] and construct validity in assessing functional independence among people with stroke, head injury and MS [40].

Table 1: Demographic characteristics of participants.

Characteristic		Centre 1 (n =36)	Centre 2 (n= 6)	Centre 3 (n=2)	Centre 4 (n=x)	Total (n=44)
Age (yr), Mean (SD) (Range)		48.1 (11.7) (21-65)	49.3 (12.7) (27-59)	53 (8.4) 47 and 59		48.5 (11.6) 21-65
Sex, number female (%)		29 (80.5)	2 (33.4)	2 (100)		33 (75)
Ethnicity n (%)	NZ European	36 (100)	5 (83.3)	2 (100)		43 (97.7)
	Māori	0	1 (16.6)	0		1 (2.3)
	American	0	0	0		0
	Others	0	0	0		0
Employment status n (%)	Employed	16 (44.4)	2 (83.3)	1 (50)		19 (43.2)
	Unemployed	20 (55.5)	4 (16.6)	1 (50)		25 (54.7)
Age at disease onset (yr), mean (SD) and range		36.1 (10.5) (19-52)	37.3 (13.3) (21-52)	38.5 (2.1) 37 and 40		36.4 (10.6) (21-52)
Disease duration (yr), mean (SD) and range		12 (8.9) (2-40)	12 (8.3) (3-26)	14.5 (10.6) (7 and 19)		12.1(10) (2-40)
Cerebellar ataxia signs n (%)	Gait ataxia	13 (36.1)	5 (83.3)	2 (100)		20 (45.5)
	Dysmetria	29 (80.5)	5 (83.3)	1 (50)		35 (79.5)
	Speech	20 (55.5)	1 (16.6)	1 (50)		22 (50)
	Nystagmus	21 (58.3)	4 (66.6)	0		25 (56.8)
Use of walking aids, n (%)	Users	12 (33.3)	1 (16.6)	1 (50)		14 (31.8)
	Non-users	24 (66.6)	5 (83.3)	1 (50)		30 (68.1)

The EDSS is a scale used to quantify disability in people with MS. There are eight functional systems that are scored using the Functional System Score (FSS). Based on the FSS scores and on ambulation, the EDSS is scored between 0 to 10 each grade rising in 0.5 increments except between the stages 0 and 1. The higher scores suggest greater disability due to MS. The EDSS demonstrates adequate inter-rater reliability (ICC=0.78), variable intra-rater reliability (ICCs between 0.62 and 0.94), and adequate convergent and discriminant validity among people with MS [41].

Current status of the study

Data collection at NZ is complete and yielded 44 participants. The data collection at the USA is in progress and will continue till the end of December 2014 or until the total sample size reached 60.

Data analysis

Non-parametric tests will be used for the ordinal scales (BBS, ICARS, SARA, Barthel Index and EDSS) and parametric test will be applied to the one interval scale (TUG). The intra class correlation coefficient (ICC) with one way random model and absolute agreement will be used to determine intra-rater and inter-rater reliability, and Cronbach alpha (α) for internal consistency. The ICC and α will be interpreted as follows: <0.50 or >-0.50 as weak, those between 0.5 and 0.79 or -0.5 and -0.79 as moderate, and those > 0.8 or <-0.8 as strong.

Spearman correlation coefficient, bivariate analysis of a non-parametric sample is used to establish constructs of validity. Correlation coefficients of <0.50 or >-0.50 will be interpreted as weak, those between 0.5 and 0.79 or -0.5 and -0.79 as moderate, and those > 0.8 or <-0.8 as strong.

Group differences of scores between users and non-users of assistive walking devices will be considered for establishing predictive validity. The Mann-Whitney U Test will be used to determine group differences for the ordinal scales, and the independent t test for the TUG, as it is an interval scale. Further, a receiver operating characteristics (ROC) curve will be constructed to determine the cut off score, sensitivity, and specificity of the measures to discriminate the use of an assistive device. In addition to determine which measure had a better predictive ability, the 'area under the curve' (AUC) of the ROC will be examined. The MCID will be estimated using a data

driven method proposed by Werwich et.al [42, 43]. The Cronbach alpha of the measures of balance will be used to estimate the Standard Error of Measurement (SEM) that reflected the MCID.

The following formula was used to determine the MCID.

$$\text{SEM or estimated MCID} = S_1 \times \sqrt{1 - r_{\alpha}}$$

S_1 = standard error at baseline, r_{α} = Cronbach's alpha coefficient for internal consistency.

The computer program "Statistical Package for the Social Sciences" (SPSS) statistics version 20, will be used to analyse the data.

Results and Discussion

The demographic data of 44 participants from NZ are presented in Table 1. Two hundred and twenty patients with MS were screened. Of these, seventy-three were eligible for the study and the remainders were excluded for the following reasons: did not meet age criteria (n=24), was wheel chair or bed bound (n=7), no ataxia (n=106), incorrect NHI ID number (n=3), moved out of NZ (n=5), or deceased (n=2). Of the seventy-three eligible, 44 gave informed consent and were recruited to the study.

The mean age of participants was 48.5 years (SD±11.6), 33 (75 %) were female, forty three (97.72%) were NZ Europeans and one (2.27%) was Māori. Nineteen participants (43.2 %) were employed and the rest were either unemployed or retired from work. The mean duration of their disease was 12.1 (SD±10) years and the mean age at disease onset was 36.4 (SD±10.6) years. Twenty-nine participants (65.9 %) did not use an assistive device for ambulation and 15 (34.1%) were assistive device users. Some of the assistive devices the participants used were: 1 stick (n=5), 1 elbow crutch (n=4), 2 elbow crutches (n=2), 1 quadripod (four-legged cane) (n=1) and rollator (rolling walker) (n=3). Four assessments were performed at participant's home and the remaining (n=40) took place at one of the research centres. EDSS assessments for eight participants were performed by a neurologist, and for the remaining 36 participants, their EDSS scores were retrieved from their medical records. Flow of data collection and the proposed scheme for psychometric property analysis can be found at Figure 1.

The findings of this study will be the first to investigate and report on reliability, validity and interpretability of four measures of balance among participants with MS having secondary cerebellar involvement. The measures of balance included are two generic scales (BBS and TUG) and two cerebellar-specific scales (PG of ICARS and SARA Bal).

Clinical Significance of the Findings

Finding evidence for the use of generic measures among people with cerebellar ataxia will be one of the key features of this study. Previous studies report the frequent use of generic measures by clinicians for the assessment of balance among people with cerebellar ataxia [20, 44]. However, the appropriateness of using these measures is unknown. The findings of this study will assist clinicians in choosing suitable generic measure that is psychometrically sound for the assessment of balance among the target population.

Secondly, this is the first study to report MCID for the measures of balance. Since this study will not re assess participants following an intervention where arguably a change in score could be expected, a data driven method will be used to establish the MCID. The MCID to be established is a reflection of the standard error of measurement (SEM) for the measures of balance and could be considered as a “proxy” for the MCID. The term “proxy” in statistics refers to a value that is probably not in itself of any great interest, but from which a variable of interest can be obtained. Previous studies have found an excellent agreement between the MCID and one SEM on Chronic Heart Disease Questionnaire tested among people with cardiovascular diseases [43]. Therefore the derived MCID, will guide clinicians and researchers to report on the treatment effects from a clinical perspective in terms of a difference in score that is perceived as a health benefit by the patient.

Lastly, the derived cut-off scores may assist clinicians in making decisions on prescribing assistive devices for walking among people with MS and secondary cerebellar involvement. Assistive device use improves the persons balance, mobility and functional ability; facilitates a generalized wellbeing and in addition, reduces the decline of functional status and the burden of health care [45-47]. Therefore prescribing assistive device at the right time becomes critical. Since we aim to determine a cut-off score with high sensitivity (90%), one may rely on the measures of balance to effectively identify people who may require an assistive device.

This study has several strengths that include, sample size adequate to fulfill ‘good’ quality research according to the COSMIN criteria, diversity of the geographic representation of participants, inclusion of neurologists to identify and assess disease severity of the participants, balance assessments performed by qualified physiotherapists, use of the ‘Standardised video recording protocol’ to standardise the assessment across the study centres. Some of the disadvantages of the proposed methods are: the influence of fatigue on participants’ performance, lack of standardizing the time of assessment that may have influenced participants’ potential. MS is a condition that affects multiple systems, and although people were specifically recruited with cerebellar ataxia, it is likely that other systems may influence the presentation of balance dysfunction.

Conclusion

The findings of this study will have high relevance in that,

recommendations on a standardized core set of measures for the assessment of balance among people with cerebellar ataxia will be proposed. In addition, this study will find evidence for the use of generic measures (BBS and TUG) among people with cerebellar ataxia and cerebellar-specific (PG of ICARS and SARA Bal) among people with MS and cerebellar ataxia. Values of high clinical relevance such as the MCID, cut-off scores, sensitivity and specificity will be reported.

Acknowledgment

The study authors would like to acknowledge the research administrator Dr. Marina Moss and five research assistants, Priya Kannan, Brooke Klatt, Abinash Babu, Jacob Jayamoorthy and Augustine Joshua who have been supporting us with data collection.

References

1. Compston A, Coles A . Multiple sclerosis. *Lancet*. 2008; 372: 1502-1517.
2. Pugliatti M, Sotgiu S, Rosati G . The worldwide prevalence of multiple sclerosis. *Clin Neurol Neurosurg*. 2002; 104: 182-191.
3. Rosati G . The prevalence of multiple sclerosis in the world: an update. *Neurol Sci*. 2001; 22: 117-139.
4. Alla S, Mason DF2 . Multiple sclerosis in New Zealand. *J Clin Neurosci*. 2014; 21: 1288-1291.
5. Taylor BV, Pearson JF, Clarke G, Mason DF, Abernethy DA, Willoughby E, et al . MS prevalence in New Zealand, an ethnically and latitudinally diverse country. *Mult Scler*. 2010; 16: 1422-1431.
6. Fox R, Hersh C. Multiple Sclerosis. Disease Management Project Chapters Neurology- Multiple Sclerosis [Internet]. 2014.
7. Mills RJ, Yap L, Young CA . Treatment for ataxia in multiple sclerosis. *Cochrane Database Syst Rev*. 2007; : CD005029.
8. Koch M. What's going wrong in ataxia and tremor in MS? *MS in focus*. 2009 (13): 3.
9. Warren K, Catz I, 2009. Cerebellar MS: A Case Study. united spinal's MS scene The larger view on multiple sclerosis.
10. Winser SJ, Hale L, Claydon LS, Smith C. Outcome measures for the assessment of balance and posture control in cerebellar ataxia. *Physical Therapy Reviews*. 2013; 18:117-133. PubMed PMID: 86141089.
11. Ataxia U. Management of the ataxias: towards best clinical practice. London: Ataxia UK. 2009.
12. Williams K, Hoang P. Ataxia and tremor in people with multiple sclerosis (MS). 2009.
13. Armutlu K. Physical Therapy and Rehabilitation Applications for Ataxic Patients. . International encyclopedia of rehabilitation. Turkey2013.
14. Martin CL, Tan D, Bragge P, Bialocerkowski A . Effectiveness of physiotherapy for adults with cerebellar dysfunction: a systematic review. *Clin Rehabil*. 2009; 23: 15-26.
15. Jacobs JV, Horak FB, Tran VK, Nutt JG . Multiple balance tests improve the assessment of postural stability in subjects with Parkinson's disease. *J Neurol Neurosurg Psychiatry*. 2006; 77: 322-326.
16. Berg K. Measuring balance in the elderly: preliminary development of an instrument. *Physiotherapy Canada*. 1989; 41:304-311.
17. Khasnis A, Gokula RM . Romberg's test. *J Postgrad Med*. 2003; 49: 169-172.
18. Martin CL, Phillips BA, Kilpatrick TJ, Butzkueven H, Tubridy N, McDonald E, et al. Gait and balance impairment in early multiple sclerosis in the absence of clinical disability. *Mult Scler*. 2006; 12: 620-628.
19. Winser SJ, Smith CM, Hale LA, Claydon LS, Whitney SL, Mehta P . Systematic review of the psychometric properties of balance measures for cerebellar ataxia. *Clin Rehabil*. 2014; .

20. Winsor SJ, Smith C, Hale LA, Claydon LS, Whitney SL . Balance outcome measures in cerebellar ataxia: a Delphi survey. *Disabil Rehabil.* 2014; .
21. Terwee CB, Mokkink LB, Knol DL, Ostelo RW, Bouter LM, de Vet HC . Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. *Qual Life Res.* 2012; 21: 651-657.
22. Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. COSMIN checklist manual 2012.
23. Jaeschke R, Singer J, Guyatt GH . Measurement of health status. Ascertaining the minimal clinically important difference. *Control Clin Trials.* 1989; 10: 407-415.
24. Berg K, Wood-Dauphinee S, Williams J. The Balance Scale: reliability assessment with elderly residents and patients with an acute stroke. *Scandinavian journal of rehabilitation medicine.* 1995; 27: 27-36.
25. Cattaneo D, Jonsdottir J, Repetti S . Reliability of four scales on balance disorders in persons with multiple sclerosis. *Disabil Rehabil.* 2007; 29: 1920-1925.
26. Learmonth YC, Paul L, McFadyen AK, Mattison P, Miller L. Reliability and clinical significance of mobility and balance assessments in multiple sclerosis. *International Journal of Rehabilitation Research.* 2012; 35: 69-74.
27. Cattaneo D, Regola A, Meotti M . Validity of six balance disorders scales in persons with multiple sclerosis. *Disabil Rehabil.* 2006; 28: 789-795.
28. Podsiadlo D, Richardson S. The timed" Up & Go": a test of basic functional mobility for frail elderly persons. *Journal of the American geriatrics Society.* 1991; 39:142-148.
29. Weyer A, Abele M, Schmitz-Hübsch T, Schoch B, Frings M, Timmann D, et al . Reliability and validity of the scale for the assessment and rating of ataxia: a study in 64 ataxia patients. *Mov Disord.* 2007; 22: 1633-1637.
30. Schmitz-Hübsch T, du Montcel ST, Baliko L, Berciano J, Boesch S, Depondt C, et al. Scale for the assessment and rating of ataxia: development of a new clinical scale. *Neurology.* 2006; 66: 1717-1720.
31. Yabe I, Matsushima M, Soma H, Basri R, Sasaki H . Usefulness of the Scale for Assessment and Rating of Ataxia (SARA). *J Neurol Sci.* 2008; 266: 164-166.
32. Bürk K, Mälzig U, Wolf S, Heck S, Dimitriadis K, Schmitz-Hübsch T, et al. Comparison of three clinical rating scales in Friedreich ataxia (FRDA). *Mov Disord.* 2009; 24: 1779-1784.
33. Lee YC, Liao YC, Wang PS, Lee IH, Lin KP, Soong BW . Comparison of cerebellar ataxias: A three-year prospective longitudinal assessment. *Mov Disord.* 2011; 26: 2081-2087.
34. Schmitz-Hübsch T, Tezenas du Montcel S, Baliko L, Boesch S, Bonato S, Fancellu R, et al. Reliability and validity of the International Cooperative Ataxia Rating Scale: a study in 156 spinocerebellar ataxia patients. *Movement disorders.* 2006; 21: 699-704.
35. Schoch B, Regel JP, Frings M, Gerwig M, Maschke M, Neuhäuser M, et al. Reliability and validity of ICARS in focal cerebellar lesions. *Mov Disord.* 2007; 22: 2162-2169.
36. Cano SJ, Hobart JC, Hart PE, Korlipara LV, Schapira AH, Cooper JM . International Cooperative Ataxia Rating Scale (ICARS): appropriate for studies of Friedreich's ataxia? *Mov Disord.* 2005; 20: 1585-1591.
37. Morton SM, Tseng YW, Zackowski KM, Daline JR, Bastian AJ . Longitudinal tracking of gait and balance impairments in cerebellar disease. *Mov Disord.* 2010; 25: 1944-1952.
38. Hsueh I-P, Lee M-M, Hsieh C-L. Psychometric characteristics of the Barthel activities of daily living index in stroke patients. *Journal-formosan medical association.* 2001; 100: 526-532.
39. Hsueh I-P, Lin J-H, Jeng J-S, Hsieh C-L. Comparison of the psychometric characteristics of the functional independence measure, 5 item Barthel index, and 10 item Barthel index in patients with stroke. *Journal of Neurology, Neurosurgery & Psychiatry.* 2002; 73:188-190.
40. Hobart JC, Thompson AJ . The five item Barthel index. *J Neurol Neurosurg Psychiatry.* 2001; 71: 225-230.
41. Hobart J, Freeman J, Thompson A . Kurtzke scales revisited: the application of psychometric methods to clinical intuition. *Brain.* 2000; 123 : 1027-1040.
42. Wywich KW, Nienaber NA, Tierney WM, Wolinsky FD . Linking clinical relevance and statistical significance in evaluating intra-individual changes in health-related quality of life. *Med Care.* 1999; 37: 469-478.
43. Wywich KW, Tierney WM, Wolinsky FD. Further evidence supporting an SEM-based criterion for identifying meaningful intra-individual changes in health-related quality of life. *Journal of clinical epidemiology.* 1999; 52:861-873.
44. Fonteyn EM, Keus SH, Verstappen CC, van de Warrenburg BP . Physiotherapy in degenerative cerebellar ataxias: utilisation, patient satisfaction, and professional expertise. *Cerebellum.* 2013; 12: 841-847.
45. Bateni H, Maki BE . Assistive devices for balance and mobility: benefits, demands, and adverse consequences. *Arch Phys Med Rehabil.* 2005; 86: 134-145.
46. Faruqi SR, Jaeblo T . Ambulatory assistive devices in orthopaedics: uses and modifications. *J Am Acad Orthop Surg.* 2010; 18: 41-50.
47. Bradley SM, Hernandez CR . Geriatric assistive devices. *Am Fam Physician.* 2011; 84: 405-411.